Analog Integrated Circuits and Signal Processing, 45, 71–78, 2005

An Adaptive Biased Single-Stage CMOS Operational Amplifier with a Novel Rail-to-Rail Constant-g_m Input Stage

MOHAMMAD MAHDI AHMADI

Department of Electrical and Computer Engineering, University of Calgary, Calgary, Canada

Received December 31, 2003; Revised July 23, 2004; Accepted August 16, 2004

Abstract. Using rail-to-rail (R-R) swing analog circuits has become almost mandatory in the design of low supply voltage circuits. In this paper, a new architecture for constant-g_m rail-to-rail input stages is presented. The design features a less than 5% deviation in g_m over the entire range of the input common-mode voltage. Furthermore, a new structure for folded cascode amplifier based on the use of a floating current source is presented. By employing these techniques, a low-power operational amplifier (op-amp) with 100 MHz unity-gain bandwidth, 106 dB gain, 60° phase margin, 2.65 V swing, and 6.4 nV/$\sqrt{}$Hz input-referred noise with rail-to-rail input common-mode range is realized in a 0.8 μm CMOS technology. This amplifier dissipates 10 mW from a 3 V power supply.

Key Words: CMOS, operational amplifier, rail-to-rail, constant-g_m

1. Introduction

Operational amplifiers are among the most widely used functional blocks for high-level analog and mixed-signal integrated circuits. The overall performance of many circuits and systems depends mainly on the quality of the op-amp used in their design. High-speed op-amps with R-R input common-mode range have a wide range of applications in high-speed and low-voltage signal processing circuits.

At large supply voltages, there is a tradeoff among speed, gain and power. Signal swing becomes yet another performance metric to be considered when designing op-amps at low supply voltages [1]. For high-swing op-amps operating at low supply voltages, output swing is not usually a problem because class-A or class-AB output stages or even folded cascode architectures can drive loads close to the power supply rails. However, designing high-swing input stages remains a challenging task.

In order to have a R-R input common-mode range, a traditional approach is to use two complementary differential pairs to form the input stage (Fig. 1) [2]. The basic idea is to have at least one of the input differential pairs active over the entire input common mode voltage, V_{CM}. However, this circuit suffers from some drawbacks. The total input transconductance, g_{mT}, is given by the sum of the transconductances of the NMOS and PMOS differential pairs. When V_{CM} is at the midway of supply voltages, both NMOS and PMOS differential pairs are active. As a result, g_{mT} could be much larger compared to when V_{CM} is near one of supply voltages. Therefore, the deviations in g_{mT} as a function of V_{CM} can be as much as 100% (Fig. 2). Furthermore, at large signal regime, the large signal output current is much larger when V_{CM} is halfway between the supply voltages. This means that the slew-rate of a conventional single-stage or two-stage amplifier with this input stage will be a function of V_{CM}. These undesirable effects complicate the frequency compensation and result in harmonic distortion [3].

Limitations of the circuit depicted in Fig. 1 provided an incentive for analog designers to create input stage circuits that have a constant input transconductance over the whole range of V_{CM} [3–10]. As will be explained in Section 3, using these input stages in single-stage op-amps results in poor tradeoff between power consumption and speed of the op-amp. Indeed, almost all of the R-R constant-g_m operational amplifiers have been realized in two- or three-stage configurations, which result in limited bandwidth and speed for the amplifier.

In this paper, a new architecture for constant-g_m R-R input stages is presented that has less than 5%

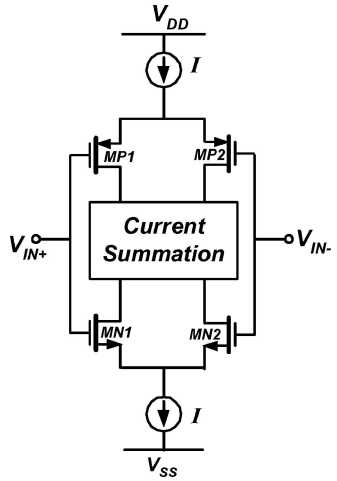

Fig. 1. Typical rail-to-rail CMOS input stage.

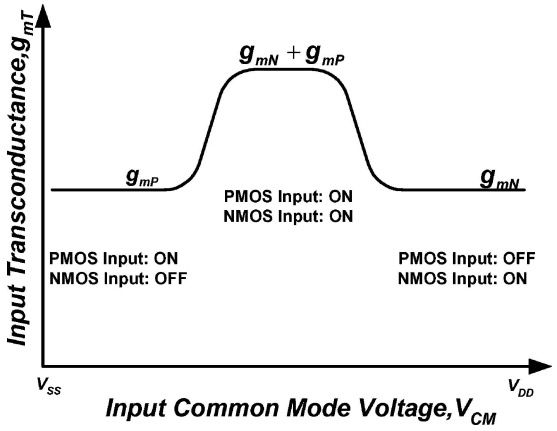

Fig. 2. g_{mT} versus V_{CM}.

deviation in g_m over the entire range of V_{CM}. Furthermore, a new structure for folded cascode amplifier is presented that allows designing a power-efficient R-R input single-stage op-amp [11]. This paper is organized as follow: input stage and its operation are explained in Section 2. In Section 3, the limitations of using R-R input stages in single-stage op-amps is described and subsequently a new architecture for folded cascode amplifier based on the use of a floating current source which mitigates those limitations is presented. In Section 4, results are presented to illustrate the effectiveness of presented approaches. Section 5 is the conclusion.

2. Rail-to-Rail Constant g_m Input Stage

Based on the MOS square law equation ($I = K(W/L)(V_{\mathrm{GS}} - V_t)^2$), if the W/L ratio of a MOS transistor is kept constant and its current is increased by a factor of 4, its g_m will be doubled. This fact is the principle for most of the constant-g_m input stages presented so far [3–8]. The principal drawback of this approach is revealed in CMOS submicron processes in which the square law equation in a MOS transistor is no longer valid and g_m is less dependent on the transistor current. That is, the current increment factor should be higher than 4, i.e. 6 or 7, to double the g_m of an MOS transistor. As a result, when V_{CM} is near power supply rails, power consumption increases considerably. This makes the design of the current summation (folded cascode) circuit complicated.

There are some other structures [9, 10], which do not operate based on the MOS square law equation. However, their input transconductance are subject to large variations. The operation of the novel architecture presented here is not based on the MOS square law equation. Besides, it maintains a nearly constant g_m in the whole range of V_{CM} and shows superior performance compared to other structures.

The idea of this technique is to activate another pair similar to the active pair when the opposite pair has lost required gate bias for proper operation. As a result, at the extreme input voltage ranges where one input pair has lost the required gate bias for proper operation, two similar pairs of other polarity generate signal current in parallel. Consequently, the input transconductance is doubled. To implement this approach, as shown in Fig. 3, three differential pairs of each polarity are used. All the PMOS differential pairs have equal W/L ratios. All the NMOS differential pair transistors are also identical.

When V_{CM} is at the midrange of supply voltages, amongst the NMOS transistors, just pairs *MN1–MN2* and *MP1–MP2* generate signal current. In this region of operation, the current from the current source I_{N3} flows through transistors *MN5* and *MN6* and sinks all the tail current of the pair *MP3–MP4*, resulting in the pair being off. On the other side, the current of the source I_{P3} flows through transistors *MP5* and *MP6* and supplies all the tail current of the pair *MN3–MP4*, resulting in *MN3–MP4* switching off. Hence, only four devices *MN1*, *MN2*, *MP1*, and *MP2* generate signal current and g_{mT} is equal to the sum of input NMOS transistor

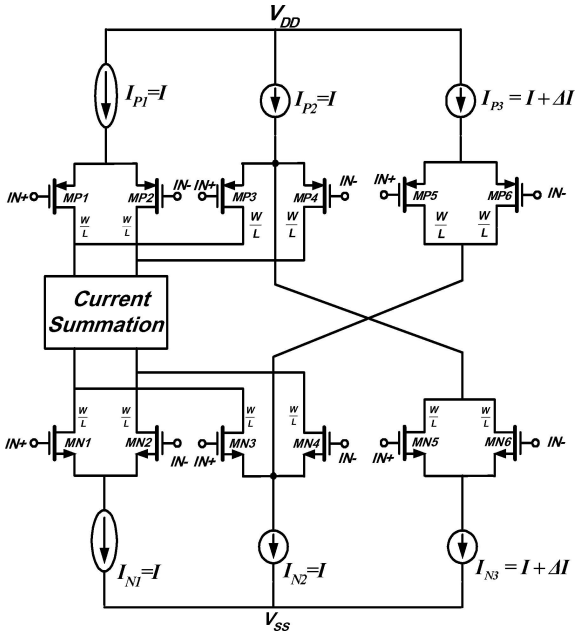

Fig. 3. New constant-g_m rail to rail input stage.

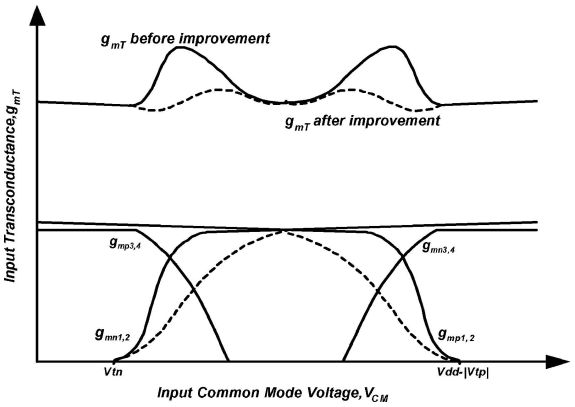

Fig. 4. The principle of reducing the g_{mT} deviations in new R-R input stage. The solid and the dotted lines are the g_{mT} and its components before and after the improvement, respectively.

transconductance, g_{mn}, and input PMOS transistor transconductance, g_{mp}.

Assuming that the nominal values of the g_m of NMOS and PMOS pairs are equal, $g_{mn} = g_{mp} = g_m$, then g_{mT} is equal to $2g_m$. Besides, there is a total current of $2I$ available for the current summer when the circuit is in the slew-limited mode.

When $V_{\rm CM}$ is near $V_{\rm DD}$, none of the PMOS transistors have a sufficient $V_{\rm GS}$ to remain active. Since there are no other paths for tail current sources I_{P1} and I_{P3}, these current sources are deactivated. Therefore, the tail current source I_{N2} is not supplied from elsewhere and flows through the pair *MN3–MN4*. Therefore, this pair turns on and works in parallel with the main pair *MN1–MN2*. As a result, g_{mT} is equal to $2g_{mn}$ which is equal to $2g_m$. There is also a total current of $2I$ available in the limiting situation.

At the other extreme point of $V_{\rm CM}$, similarly, I_{N3} is pushed to the cutoff state and I_{P2} flows through *MP3* and *MP4*. Consequently, the equivalent transconductance is $2g_{mp} = 2g_{m.}$ There is also a total current of $2I$ available in the slew-limited regime.

If due to transistor mismatches, I_{N2} and I_{P2} exceed I slightly, or I_{N3} and I_{P3} become slightly smaller than I, the pairs *MN3–MN4* and *MP3–MP4* will turn on when $V_{\rm CM}$ is in the midway of supply voltages. This undesired activation of pairs *MN3–MN4* and *MP3–MP4* increases g_{mT} and enhances the g_{mT} deviations. By setting the value of current sources I_{P3} and I_{N3} slightly higher than I, i.e. $I + \Delta I$, this problem can be eliminated.

In this circuit, by handling the W/L ratio and $V_{\rm GS}$ of the transistors realizing the tail current sources, the g_{mT} deviation can be further reduced. The concept of decreasing the g_{mT} deviation is shown in Fig. 4. The solid and the dotted lines are the g_{mT} and its components before and after the improvement, respectively. The overhangs in the g_{mT} curve are because of the large slope in the g_m of differential pairs in takeover regions. As illustrated in Fig. 4, as the slopes of the curves are reduced, the g_{mT} curve becomes smoother. To reduce this slope, i.e. the slope of the g_m-$V_{\rm CM}$ curve of a differential pair in takeover regions, the effective voltage, $V_{\rm DS,sat}$, of the tail current source transistor should be increased. That is, the effective voltages of the transistors which realize the current sources I_{N1} and I_{P1} should be increased. This can be done simply by reducing their W/L ratios and increasing their $V_{\rm GS}$.

By cascoding the tail current sources I_{N2}, I_{N3}, I_{P2}, and I_{P3}, the deviations in g_{mT} are further reduced. The post-layout simulation result of the g_m of the input stage is shown in Fig. 5. With this topology, the g_{mT} deviations could be less than 5%, however, since six differential pairs have been used to realize the R-R input stage, the input capacitance of this op-amp is somewhat considerable, which may not be acceptable in some applications. On the other hand, this large input capacitance results in smaller

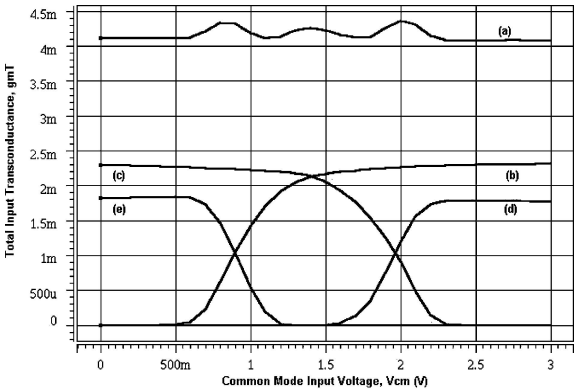

Fig. 5. g_{mT} and its components versus V_{CM} , (a) g_{mT} , (b) $g_{mn1,2}$ (c) $g_{mp1,2}$, (d) $g_{mn3,4}$ and (e) $g_{mp3,4}$.

input referred noise, which is desirable in many applications.

3. Adaptive Biased Current Summation Circuit

Another important block in an op-amp with R-R input common-mode range is the current summing circuit. The conventional approach for realizing this circuit is shown in Fig. 6. However, when this circuit is used in single-stage amplifiers, some problems are caused in the frequency compensation of the amplifier. These problems can be addressed as follows with reference to Fig. 6:

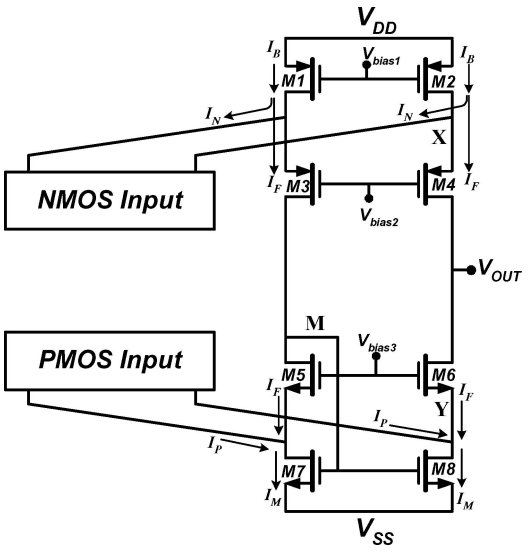

Fig. 6. Conventional method for biasing current summation circuit.

The current of NMOS input devices, I_N, changes between zero for $V_{\mathrm{CM}} = V_{\mathrm{SS}}$ and more than $2I_n$ for $V_{\mathrm{CM}} = V_{\mathrm{DD}}$; where I_n is the NMOS input stage bias current for when V_{CM} is at midrange of supply voltages. As a result, the bias current of $M1$ and $M2$, I_B, not only should be able to supply the bias current I_F of transistors $M3$ to $M6$, but also should be high enough to handle the large changes in the current of NMOS input devices (from zero to $2I_n$).

Note that when V_{CM} is at the midway of supply voltages or close to V_{SS}, I_N is reduced, resulting in a considerable increases in I_F. As a result, the transconductance and output resistance of transistors $M3$ to $M8$ change considerably, leading to large variations in pole-zero locations and also the open-loop gain of the op-amp.

These large variations in pole-zero locations complicate the frequency compensation and prevent the optimum usage of power in order to enhance the bandwidth. Here we can have a glance at the pole-zero locations and gain of this amplifier. The poles and zeros of the op-amp depicted in Fig. 6 are located at following frequencies:

$$P_1 = -\frac{1}{C_{\mathrm{out}} R_{\mathrm{out}}} \quad P_2 = -\frac{g_{m4}}{C_x} \quad P_3 = -\frac{g_{m6}}{C_y}$$

$$P_{4,5} = -\frac{g_{m5}}{2C_y}\left[1 + \sqrt{1 - \frac{4g_{m7}C_y}{g_{m5}C_M}}\right]$$

$$Z_1 = -\frac{g_{m4}g_{m6}(g_{mn} + g_{\mathrm{mp}})}{g_{mn}g_{m4}C_y + g_{\mathrm{mp}}g_{m6}C_x}$$

$$Z_{2,3} = -\frac{g_{m5}}{2C_y}\left[1 \pm \sqrt{1 - \frac{8g_{m7}C_Y}{g_{m5}C_M}}\right]$$

where C_X, C_Y, C_M and C_{OUT} denote the total capacitances at nodes X, Y, M and OUT respectively. g_{mi}, g_{mn} and g_{mp} denote the transconductances of transistor Mi, the NMOS input stage and the PMOS input stage, respectively, and R_{out} is the output resistance of the op-amp obtained from:

$$R_{\mathrm{OUT}} = (g_{m4}r_{O4}(r_{O2} \parallel r_{\mathrm{ON}})) \parallel (g_{m6}r_{O6}(r_{O8} \parallel r_{\mathrm{OP}}))$$

where r_{oi} is the output resistance of transistor Mi. As it is obvious in the above equations, any variations in the bias current of the transistors considerably change the frequency response and output resistance as well as the open-loop gain of the op-amp.

These variations in the bias current of the current summation transistors are not very important in

two- or multi-stage amplifiers, which use R-R input stage. This is mainly due to the fact that in two- or multi-stage amplifiers the voltage swing at the output node of the first stage is not a restrictive factor. Therefore, the W/L ratio of the output transistors of the first stage should not be high. This results in small parasitic capacitances at nodes X, Y, and M. Consequently, the poles located at these nodes are at fairly high frequencies and their effects can be safely ignored. Indeed, in these amplifiers the bandwidth is restricted to the poles of the second or other stages, which inherently are at lower frequencies compare to the poles located at nodes X, Y and M.

However, for single-stage amplifiers, the output node of the current summation circuit is the output node of the op-amp. As a result, the voltage swing at this node determines the op-amp output voltage swing. Hence, the W/L ratio of the devices $M1$ to $M8$ should be high, resulting in high parasitic capacitances at nodes X and Y. Therefore, the small-signal parameters of these devices affect the frequency response of the amplifier considerably. In order to optimize the power consumption, the variations in the pole-zero locations due to the bias current variations should be minimized as much as possible.

From detailed analysis, it becomes clear that the location of the poles, P_2, P_3, $P_{4,5}$ and zeros, Z_1, $Z_{2,3}$ is directly related to the transconductance of transistors $M3$–$M6$. In fact, the transconductance and output resistance of these transistors have the most important role in the pole-zero locations and gain of the op-amp. Therefore, stabilizing the quiescent current of these transistors helps in optimizing the frequency compensation and reducing the harmonic distortion of the amplifier.

For this reason, both NMOS and PMOS cascoded transistors are used as current mirrors. Note that the current value of the current mirror changes adaptively by the changes in the input differential pair currents. The remained problem is to bias the current summation circuit.

The first approach is to use two independent current sources as depicted in Fig. 7. The principal drawback of this approach is that the noise of bias current sources of the current mirrors directly contributes to the noise of the amplifier. This is due to the unity current gain between bias current sources and the drain currents of the input transistors. Any mismatch in the bias current sources will also increase the offset of the amplifier [6].

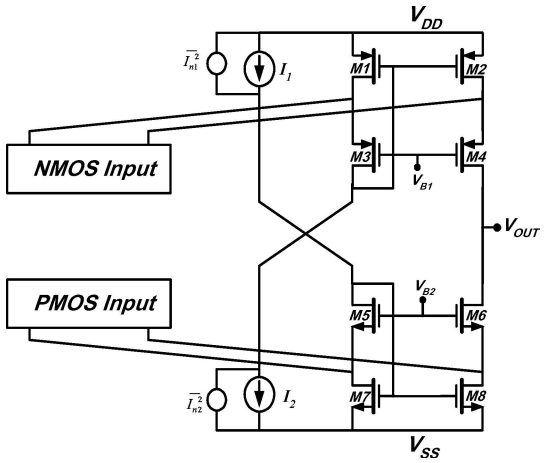

Fig. 7. A simple approach for biasing current summation circuit.

Using a floating current source between the drains of transistors $M3$ and $M5$, as shown in Fig. 8, can alleviate the mentioned problem. Besides, the bias current of transistors $M3$–$M6$ stays nearly constant. The circuit realization of the floating current source with the complete diagram of the entire amplifier is shown in Fig. 9. The value of this floating current source is determined by the MOS translinear loops $M7$, $M9$, $M12$, $M11$ and $M1$, $M10$, $M13$, $M14$. The floating current source has been used for the design of class AB amplifiers previously [6], however, in this paper, it has been employed to bias the current summation circuit.

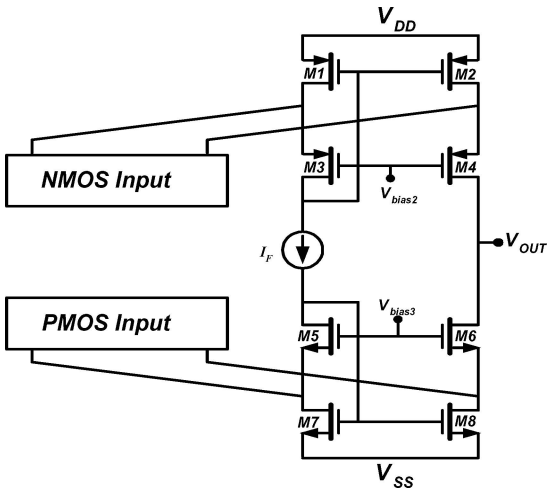

Fig. 8. Using a floating current source for biasing the current summation circuit.

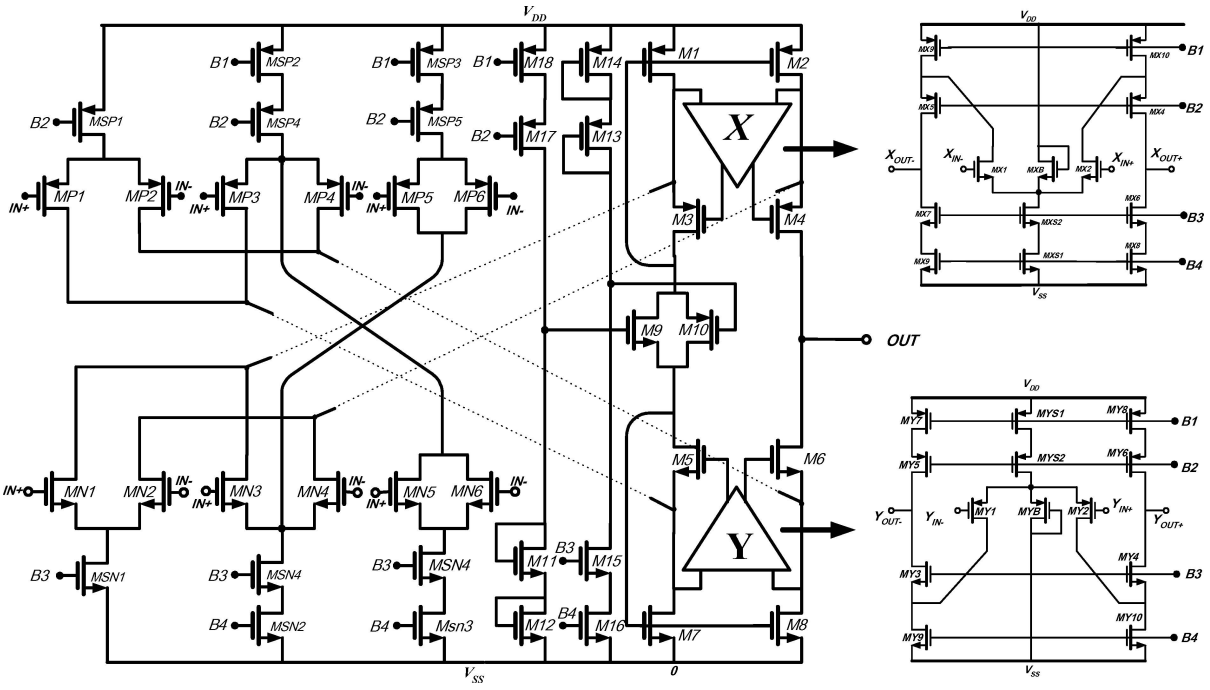

Fig. 9. Complete schematic of the amplifier: the floating current source is realized by transistors *M9* to *M18*. There are two gain-boosting amps, X and Y. The circuit realizations of the amplifiers X and Y are shown in right hand of the figure.

Gain boosting has also been employed to ensure sufficient gain for the amplifier. Indeed, by two auxiliary folded cascode amplifiers, the output resistance of the amplifier increases considerably without any degradation of frequency response [12]. The main parameter in designing the auxiliary amplifiers is their unity-gain bandwidth which should be about half of the frequency of the folding poles. That is, the unity-gain bandwidth of the auxiliary amplifiers X and Y should be chosen about half of the frequency of the poles P_2 and P_3, respectively [13].

4. Amplifier Specifications

The proposed single-stage amplifier with constant-g_m R-R input stage has been fabricated in a 0.8 μm double-poly, double-metal CMOS process provided by AMS Corporation. The core of the op-amp is shown in Fig. 10. This amplifier was used as a block in the design of a novel high frequency sampled-data filter, which operates based on neural network algorithm. The occupied die area is about 0.1 mm^2.

The experimental results show that this amplifier consumes a total power of 10 mW from a 3-V supply, while its input offset voltage is about 10 mV. The measured 3 dB bandwidth in the non-inverting unity-gain feedback configuration is about 100 MHz. Input voltage swing is from negative rail to positive rail, while the output swing is 2.65 V.

The results of post-layout simulations obtained from rigorous HSPICE and Cadence simulators revealed that the $g_m$of the input stage, as shown in Fig. 4, has a deviation of less than 5% over the entire range of V_{CM}.

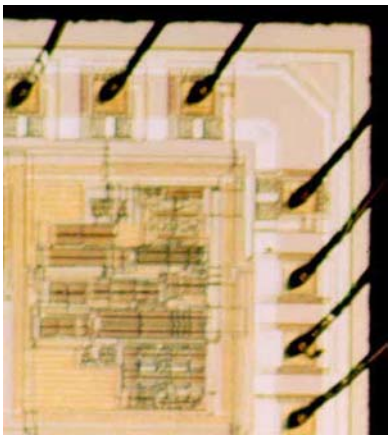

Fig. 10. Chip micrograph of the amplifier.

Table 1. Op-amp characteristics with a 3-V Supply voltage and 5pF capacitive load.

Parameter	Value
Die area	0.1 mm^2
Power Consumption	10mW
Input voltage swing	R-R
Output voltage swing	2.65V
3 dB Bandwidth in noninverting unity gain feedback	100 MHz
Input offset voltage	10 mV
Gain*	102 dB
Phase margin*	60°
Settling time 0.1*	15 nS
Settling time 0.01*	30 nS
Slew rate*	150 V/μS
CMRR @ 1 KHz*	121 dB
Input noise voltage*	6.4 nV/$\sqrt{}$Hz
g_m variation*	5%

*Simulated results

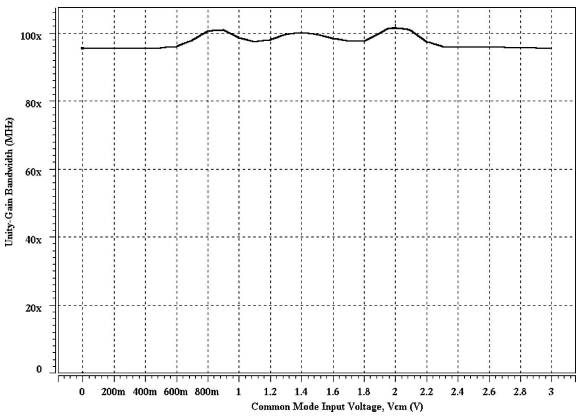

Fig. 11. Unity-gain bandwidth deviations with respect to V_{CM}.

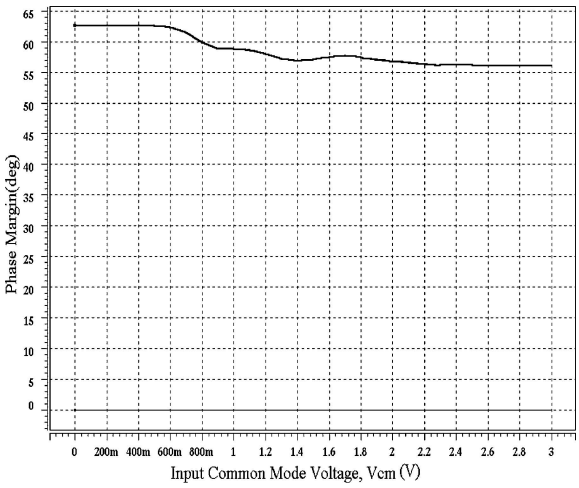

Fig. 12. Phase margin deviations with respect to V_{CM}.

The variations of the unity-gain bandwidth and phase margin versus V_{CM} are shown in Figs. 11 and 12, respectively. These figures reveal that frequency response is robust enough in terms of the changes in V_{CM}. Figure 13 shows the simulated frequency response of the op-amp with a 5 pF capacitive load when $V_{CM} = 1.5$ V. It shows that the unity-gain frequency is about 100 MHz (which is in agreement with measured results).

Figure 14 shows the step response of the amplifier in unity gain feedback configuration. The 0.1% and 0.01% settling times of the amplifier are 15 ns and 30 ns, respectively. Besides, the slew-rate of the amplifier is

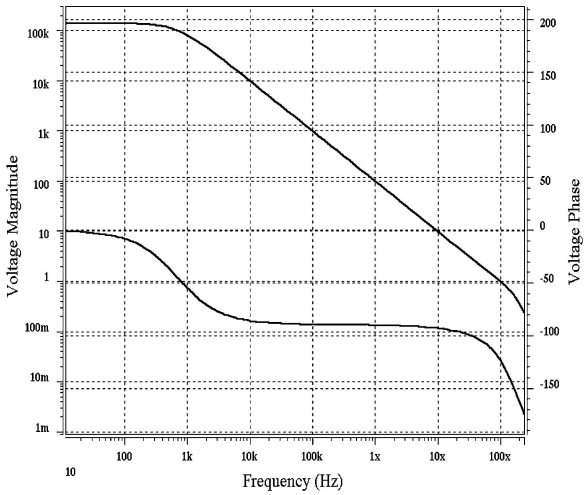

Fig. 13. Frequency response of the op-amp, $V_{CM} = 1.5$ V, $C_L = 1.5$ pF.

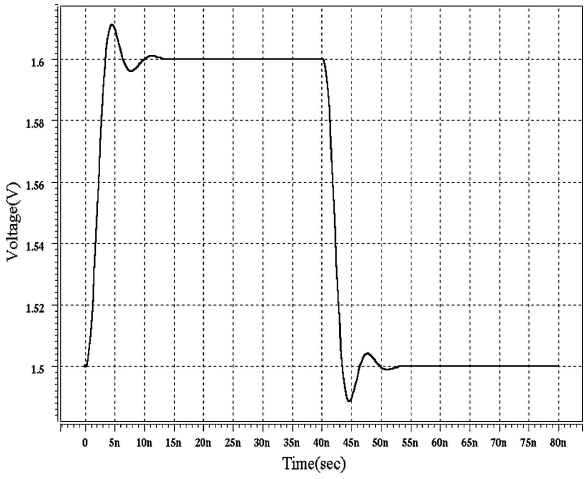

Fig. 14. Small signal step response of the op-amp.

150 V/μS, and the input referred noise is 6.4nV/$\sqrt{}$Hz. The amplifier characteristics are summarized in Table 1.

5. Conclusion

In this paper, a new architecture for R-R constant-g_m input stages is presented. The architecture is not based on the MOS square law equation, and has superior performance compared to other R-R architectures. Additionally, based on the use of a floating current source, a new architecture is presented for designing single-stage operational transconductance amplifiers. The floating current source biases the output transistors of the amplifier with minimum changes in the pole-zero locus. Employing these new architectures, a low power, high swing, high speed, and high gain single-stage operational transconductance amplifier is designed and implemented in a 0.8 μm double poly double metal CMOS process which consumes less than 10mW from a 3-V supply.

Acknowledgment

The author is grateful to reviewers for their valuable comments.

References

1. K. Gulati and H.-S. Lee, "A high-swing CMOS telescopic operational amplifier." *IEEE J. of Solid-State Circuits*, vol. 33, no. 12, pp. 2010–2019, 1998.
2. J.N. Babanezhad, "A rail-to-rail CMOS op amp." *IEEE J. of Solid-State Circuits*, vol. 23, no. 6, pp. 1414–1417, 1988.
3. J.H. Huijsing, R. Hogervorst, and K. de Langen, "Low-power low-voltage VLSI operational amplifier cells." *IEEE Trans. Circuits and Systems II*, vol. 42, no. 11, pp. 841–852, 1995.
4. M. Ryat, "Rail to rail operational transconductance amplifier." U.S. Patent 5, 208, 552, May 1993.
5. S. Sakurai and M. Ismail, "Robust design of rail-to-rail CMOS operational amplifiers for a low power supply voltage." *IEEE J. of Solid-State Circuits*, vol. 31, no. 2, pp. 146–156, 1996.
6. R. Hogervorst, J.P. Tero, R.G.H. Eschauzier, and J.H. Huijsing, "A compact power-efficient 3 V CMOS rail-to-rail input/output operational amplifier for VLSI cell libraries." *IEEE J. of Solid-State Circuits*, vol. 29, no. 12, pp. 1505–1513, 1994.
7. G. Ferri, W. Sansen, and V. Peluso, "A low-voltage fully differential constant-Gm rail-to-rail CMOS operational amplifier." *Analog Integrated Circuits and Signal Processing*, vol. 16, pp. 5–15, 1998.
8. H.-H. Lin, T. Pimenta, and M. Ismail, "Robust design of LV/LP low-distortion CMOS rail-to-rail input stages." *Analog Integrated Circuits and Signal Processing*, vol. 21, 153–162, 1999.
9. R. Hogervorst, S.M. Safai, J.P. Tero, and J.H. Huijsing, "A programmable 3-V CMOS rail-to-rail Op-amp with gain boosting for driving heavy resistive loads." In *Proceedings ISCAS*, Seattle, USA, 1995, vol. 2, pp. 1544–1547.
10. W. Redman-White, "A high bandwidth constant-g_m and slew-rate rail-to-rail CMOS input circuit and its application to analog cells for low voltage VLSI systems." *IEEE J. of Solid-State Circuits*, vol. 32, no. 5, pp. 701–711, 1997.
11. M.M. Ahmadi, R. Lotfi, and M. Sharif-Bakhtiar, "A new architecture for rail-to-rail input constant-g_m CMOS operational transconductance amplifiers." In *Proceeding* of ACM/IEEE *ISLPED*, Seoul, Korea, 2003, pp. 353–358.
12. K. Bult and G.J.G.M. Geelen, "A fast-settling CMOS op-amp for SC circuits with 90-dB DC gain." *IEEE J. of Solid-State Circuits*, vol. 25, no. 6, pp. 1379–1383, 1990.
13. M.M. Ahmadi, "A novel modeling and optimization of gain-boosted cascode amplifiers for high-speed applications." In *Proceeding* of 10th IEEE *ICECS*, Sharjah, UAE, 2003, pp. 683–686.

Mohammad Mahdi Ahmadi was born in Khansar, Iran, in 1978. He received the B.Sc. in Biomedical Engineering from Shahid Beheshti University, Tehran/Iran, in 2000 (with highest honor) and M.Sc. in Electronics Engineering from Sharif University of Technology, Tehran/Iran, in 2002 (with honor). He is currently doing his Ph.D. at the Department of Electrical and Computer Engineering, University of Calgary, Calgary/Canada.

His research interest is analog integrated circuit design for biomedical and communication applications.